

Universität Augsburg

Institut für
Mathematik

Ali Ünlü, Anatol Sargin

Mosaics for Visualizing Knowledge Structures

Preprint Nr. 31/2008 — 12. November 2008

Institut für Mathematik, Universitätsstraße, D-86135 Augsburg

<http://www.math.uni-augsburg.de/>

Impressum:

Herausgeber:

Institut für Mathematik

Universität Augsburg

86135 Augsburg

<http://www.math.uni-augsburg.de/forschung/preprint/>

ViSdP:

Ali Ünlü

Institut für Mathematik

Universität Augsburg

86135 Augsburg

Preprint: Sämtliche Rechte verbleiben den Autoren © 2008

Mosaics for Visualizing Knowledge Structures

Ali Ünlü and Anatol Sargin

Department of Computer Oriented Statistics and Data Analysis, Institute of Mathematics, University of Augsburg, D-86135 Augsburg, Germany
`ali.uenlue@math.uni-augsburg.de`, `anatol.sargin@math.uni-augsburg.de`

Summary. Mosaic plots are state-of-the-art graphics for multivariate categorical data. Knowledge structures are mathematical models that belong to the theory of knowledge spaces in psychometrics. This paper presents an application of mosaic plots to psychometric data arising from underlying knowledge structure models. In simulation trials, the scope of this graphing method in knowledge space theory is investigated.

Key words: Mosaic plot, Visualization, Knowledge structure, Knowledge space theory, Psychometrics.

1 Introduction

Graphics are a powerful tool for presenting and exploring data, especially when implemented in interactive visualization software. Graphics help in understanding data and in determining structure. They are easy to create, convenient to use, and they can present information effectively. Interactive graphical approaches become indispensable in particular when analyzing large and complex datasets. Classical books on graphics are Chambers, Cleveland, Kleiner, and Tukey (1983) and Cleveland (1994). Wilkinson (2005) defines a formal framework for constructing graphics. Unwin, Theus, and Hofmann (2006) discuss graphics for large datasets. Theus and Urbanek (2008) present interactive graphics. Cook and Swayne (2007) address dynamic graphics. Other recent books on graphics are Murrell (2005) and Young, Valero-Mora, and Friendly (2006). The ‘Handbook of Data Visualization’ by Chen, Härdle, and Unwin (2008) provides an overview of the current state of affairs.

In the graphics literature, various types of plots for categorical data have been proposed. Graphics for univariate categorical data are barcharts and spineplots (Hummel, 1996). For multivariate categorical data, the modern graphic is the mosaic plot, which was originally introduced by Hartigan and Kleiner (1981). The classical mosaic plot has been further developed and now

comes with the variations equal binsize mode and doubledecker plot (for comparing highlighted proportions), fluctuation diagram (for identifying common combinations), and multiple barcharts (for comparing conditional distributions); see Friendly (1994), Hofmann (1998, 2000, 2003, 2008), Hofmann, Siebes, and Wilhelm (2000), and Hofmann and Wilhelm (2001).

Doignon and Falmagne (1985) introduced knowledge space theory (KST). Most of the theory of knowledge spaces is presented in a monograph entitled ‘Knowledge Spaces’ by Doignon and Falmagne (1999); see also Doignon and Falmagne (1987), Falmagne (1989), and Falmagne, Koppen, Villano, Doignon, and Johannesen (1990). A comprehensive bibliography on KST, including many references on applications of KST, can be retrieved from <http://wundt.kfunigraz.ac.at/kst.php>. For concrete application examples, see in particular Albert and Lukas (1999). The theory of knowledge spaces has been successfully applied for the computerized, adaptive assessment and training of knowledge; for instance, see the Assessment and LEarning in Knowledge Spaces (ALEKS) system, a fully automated math tutor on the Internet, <http://www.aleks.com>.

In KST, (interactive) statistical graphics have not been considered so far. Prerequisite (or precedence) and Hasse diagrams are seen and used graphics. However, they do not really provide new insights. Their shape is determined by the models, they do not graphically display the raw data, and they are merely utilized for static presentation rather than exploration. In this paper, the mosaic plot is applied to response data arising from underlying knowledge structure models in KST. Mosaic plots can display the data and knowledge structure jointly in a single graphic. Aggregating over items in the mosaic plot, as is implemented in an interactive manner in visualization software, allows visually displaying the traces of the data and knowledge structure on subsets of the item set. Considering projections of knowledge structures on subsets of item sets, in turn, is important in KST (e.g., Falmagne, 2008). In simulation trials, using the basic local independence model as the data generating model, the scope of this graphing method in visually detecting the underlying knowledge structure is investigated. It is argued that mosaic plots, albeit they can become quickly intractable (space and display resolution), provide an interesting new perspective on visualizing and deriving knowledge structures from data. The results gained from visually inspecting mosaic plots, for instance, can be consulted as a reference against which to compare the findings obtained from KST numerical data analysis methods.

The graphics in this paper are created using Mondrian, a statistical data visualization software featuring modern interactive visualization techniques. Mondrian has been developed by Theus (2002) (for a review, see Ünlü and Sargin, 2008a) in the programming language Java. It is available at no cost for Windows, Mac OS X, and UNIX, by download from <http://www.rosuda.org/Mondrian/>. A first static implementation of mosaic plots in R (<http://www.r-project.org/>) was by Emerson (1998), other fully interactive implementations are Manet (Unwin, Hawkins, Hofmann, and

Siegl, 1997), Klimt (Urbanek, 2002), and the R package iplots (Urbanek and Theus, 2003).

2 Mosaic Plots

The mosaic plot was originally introduced by Hartigan and Kleiner (1981), for visualizing contingency tables. Subsequent works on mosaic plots include Friendly (1994, 1995, 1999, 2000), Hartigan and Kleiner (1984), and Hofmann (2003, 2008). This section gives a brief introduction to mosaic plots.

A mosaic plot consists of groups of rectangular tiles. Each tile corresponds to one cell of a contingency table. The tile's area is proportional to the cell's count. Its shape and location in the display are determined by how the mosaic plot is constructed. The following step-wise construction of a p -dimensional mosaic plot (with minor modification of notation) is taken from Hofmann (2008, Section 13.1).

Let X_1, X_2, \dots, X_p be p categorical variables. Let c_i be the number of categories of variable X_i ($1 \leq i \leq p$).

1. For $i = 0$, start with an initial single rectangle r_0 , of width w_0 and height h_0 . Let $i = 1$.
2. For each rectangle $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$ ($1 \leq j_l \leq c_l$ for $l = 1, 2, \dots, i-1$) proceed as follows ($r_0^{j_1 j_2 \dots j_0} := r_0$). Cut rectangle $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$ into c_i pieces: Find all observations corresponding to rectangle $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$ (for r_0 take all the observations in the data). Determine the (relative) breakdown for the variable X_i ; that is, among the observations corresponding to $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$, count the number of those observations that fall into each of the categories of X_i . Split the width (height) of rectangle $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$ into c_i pieces, where the widths (heights) are proportional to the breakdown, and keep the height (width) of each the same as $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$. Call these new rectangles $r_i^{j_1 j_2 \dots j_i}$ ($1 \leq j_l \leq c_l$ for $l = 1, 2, \dots, i$).
3. Increase i by 1.
4. While $i \leq p$, repeat steps 2 and 3 for all $r_{i-1}^{j_1 j_2 \dots j_{i-1}}$ ($1 \leq j_l \leq c_l$ for $l = 1, 2, \dots, i-1$).

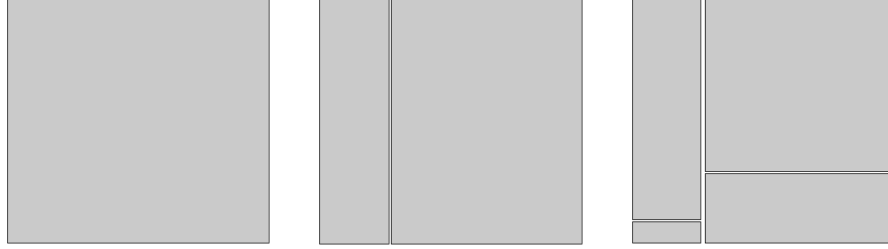
Table 1 shows a cross-tabulation of 340 German students answering two dichotomous questions on mathematical literacy (see Section 5 for details).

Figure 1 illustrates the step-wise construction of the corresponding two-dimensional mosaic plot.

This is the classical form of the mosaic plot. There are the variations equal binsize mode (each cell is allocated the same amount of space, and the information is reduced to the binary case of whether a cell is or is not empty), doubledecker plot (instead of alternately splitting the x and y axes, only the x axis is used), fluctuation diagram (each cell is allocated the same amount of space, and the cell with the maximum frequency fills its space completely,

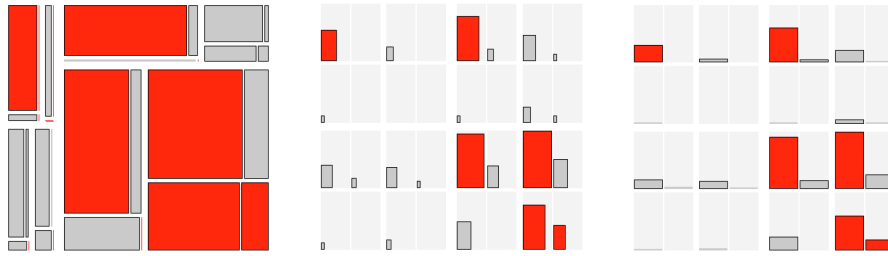
Table 1. Cross-tabulation of 340 German students according to two test items of a mathematical literacy test (see Section 5)

		Item 4		
		0	1	Totals
Item 2	0	83	8	91
	1	178	71	249
Totals		261	79	340

**Fig. 1.** Step-wise construction of two-dimensional mosaic plot of the two test items of the mathematical literacy test. Left plot: $i = 0$ and $(r_0, 340)$. Middle plot: $i = 1$, and from left to right, $(r_1^1, 91)$ and $(r_1^2, 249)$. Right plot: $i = 2$, and from left to right starting at the top, $(r_2^{11}, 83)$, $(r_2^{12}, 8)$, $(r_2^{21}, 178)$, and $(r_2^{22}, 71)$.

thus fixing the scale for the rest of the diagram), multiple barcharts (each cell is allocated the same amount of space, and only the heights of the bars in the cells are scaled).

Figure 2 shows a classical mosaic plot, fluctuation diagram, and multiple barcharts of the 340 German students answering five dichotomous test items of the mathematical literacy test, with the underlying (true) knowledge states highlighted (see Section 5 for details).

**Fig. 2.** Classical mosaic plot (left), fluctuation diagram (middle), and multiple barcharts (right) for the five test items of the mathematical literacy test dataset. Underlying knowledge states are highlighted. Details on these data and concepts are discussed in Section 5.

Experimentation with the mosaic plot (not reported in this paper) for visually detecting knowledge states from data has shown that the classical form of the mosaic plot is not as appropriate as the other two variants. This is also indicated by Figure 2. Furthermore, since it is easier to compare heights (in multiple barcharts) than areas (in fluctuation diagram), in the sequel, we confine ourselves to using multiple barcharts.

Theoretically, mosaic plots can be applied to any number of categorical variables. Practically, however, space and display resolution are limiting factors. Experimentation with the mosaic plot (not reported in this paper) has shown that, up to 11–12 dichotomous items, the fluctuation diagram and multiple barcharts are still applicable. The classical mosaic plot can be used up to 13–14 dichotomous items, albeit this variant, as compared to the former two, by far is not that effective for inspecting knowledge structures in response data.

3 Knowledge Space Theory

This section starts with a motivating small example (Falmagne, Cosyn, Doignon, and Thiéry, 2006), and then reviews some of the basic deterministic and probabilistic concepts of KST. For details, the reader is referred to the aforementioned references.

3.1 Example: Elementary Algebra

A natural starting point for a theory of knowledge assessment and training stems from the observation that some pieces of knowledge may imply other pieces of knowledge. In the context of the following example, the mastery of some algebra problem may imply the mastery of other algebra problems. Such implications between pieces of knowledge can be used to design efficient computer-based, adaptive knowledge assessment and training procedures (Falmagne and Doignon, 1988a,b, see also the ALEKS system).

As an example, we consider six (dichotomous) problems in Elementary Algebra.

- a. A car travels on the freeway at an average speed of 52 miles per hour. How many miles does it travel in 5 hours and 30 minutes?
- b. Using the pencil, mark the point at the coordinates $(1, 3)$.
- c. Perform the multiplication $4x^4y^4 \cdot 2x \cdot 5y^2$ and simplify your answer as much as possible.
- d. Find the greatest common factor of the expressions $14t^6y$ and $4tu^5y^8$. Simplify your answer as much as possible.
- e. Graph the line with slope -7 passing through the point $(-3, -2)$.
- f. Write an equation for the line that passes through the point $(-5, 3)$ and is perpendicular to the line $8x + 5y = 11$.

A plausible prerequisite (or precedence) diagram of mastery dependencies for the six Elementary Algebra problems may look like in Figure 3.

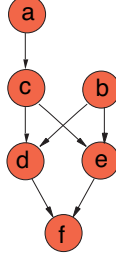


Fig. 3. Prerequisite diagram of mastery dependencies for the six Elementary Algebra problems a – f . Reflexivity and transitivity are assumed to hold and are not explicitly depicted. The mastery of Problem b is, for instance, a prerequisite for the mastery of Problem e . That is, the mastery of Problem e implies that of Problem b .

The prerequisite diagram in Figure 3 completely specifies the feasible knowledge states. A respondent can certainly master just Problem a . This does not imply mastery of any other problem. In that case, the knowledge state is $\{a\}$. However, if the respondent masters e , for instance, then a , b , and c must also be mastered. This gives the knowledge state $\{a, b, c, e\}$. Analyzing the prerequisite diagram in this way, we see that there are exactly 10 knowledge states consistent with the diagram:

$$\mathcal{K} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, b, c\}, \\ \{a, b, c, d\}, \{a, b, c, e\}, \{a, b, c, d, e\}, \{a, b, c, d, e, f\}\}.$$

This set \mathcal{K} of all possible knowledge states is called knowledge structure. These notions are formalized mathematically in the next section.

3.2 Basic Deterministic Concepts

A general concept is that of a knowledge structure.

Definition 1. A knowledge structure is a pair (Q, \mathcal{K}) in which Q is a non-empty set, and \mathcal{K} is a collection of subsets of Q containing at least the empty set \emptyset and Q . The set Q is called the domain of the knowledge structure. The elements $q \in Q$ and $K \in \mathcal{K}$ are referred to as (test) items and (knowledge) states, respectively.

As an example knowledge structure consider the one described in Section 3.1, on the domain of the six Elementary Algebra problems.

Note that this example knowledge structure is closed under union and intersection.

Definition 2. A knowledge structure (Q, \mathcal{K}) is called a knowledge space if \mathcal{K} is closed under union—that is, for all $\mathcal{F} \subset \mathcal{K}$, $\bigcup \mathcal{F} \in \mathcal{K}$. If a knowledge space (Q, \mathcal{K}) is closed under intersection—that is, for all $\mathcal{F} \subset \mathcal{K}$, $\bigcap \mathcal{F} \in \mathcal{K}$ —it is called a quasi ordinal knowledge space.

The notions of a knowledge structure and (quasi ordinal) knowledge space are at the level of persons (representing collections of knowledge states of individuals). There is another important notion, that of a surmise relation, which is at the level of items (representing collections of mastery dependencies between items).

Definition 3. Let Q be a non-empty set of items. Any quasi order—that is, reflexive and transitive binary relation—on Q is called a surmise relation.

As an example surmise relation consider the relation \preceq which corresponds to the prerequisite diagram of mastery dependencies in Figure 3:

$$\preceq = \Delta \cup \{(a, c), (a, d), (a, e), (a, f), (b, d), (b, e), (b, f), \\ (c, d), (c, e), (c, f), (d, f), (e, f)\},$$

where $\Delta = \{(x, x) : x \in \{a, b, c, d, e, f\}\}$ denotes the reflexive pairs.

Birkhoff (1937)’s theorem (see also Doignon and Falmagne, 1999, Theorem 1.49) provides a linkage between quasi ordinal knowledge spaces and surmise relations on an item set.

Theorem 1. There exists a one-to-one correspondence between the collection of all quasi ordinal knowledge spaces \mathcal{K} on a domain Q , and the collection of all surmise relations \preceq on Q . Such a correspondence is defined through the two equivalences ($p, q \in Q$, $K \subset Q$):

$$\begin{aligned} p \preceq q &: \iff [\forall K \in \mathcal{K} : \{q \in K \implies p \in K\}], \\ K \in \mathcal{K} &: \iff [\forall (p \preceq q) : \{q \in K \implies p \in K\}]. \end{aligned}$$

This theorem is important from a practical point of view. Though the quasi ordinal knowledge space and surmise relation models are empirically interpreted at two different levels, at the levels of persons and items, respectively, they are connected with each other mathematically, through Birkhoff’s theorem.

The 10 knowledge states (of the example knowledge structure) consistent with the prerequisite diagram are obtained applying the second equivalence of Birkhoff’s theorem.

3.3 Basic Probabilistic Concepts

Knowledge states are latent and not directly observable, due to random response errors. A person who is actually unable to solve an item, but does so,

makes a lucky guess. On the other hand, a person makes a careless error, if he fails to solve an item which he is capable of mastering. A probabilistic extension of the knowledge structure model covering random response errors is the basic local independence model. We use this probability model for simulating data.

Definition 4. A quadruple (Q, \mathcal{K}, p, r) is called a basic local independence model (BLIM) if

1. (Q, \mathcal{K}) is a knowledge structure, where Q is finite;
2. p is a probability distribution on \mathcal{K} —that is, $p : \mathcal{K} \rightarrow]0, 1[, K \mapsto p(K)$, with $p(K) > 0$ for any $K \in \mathcal{K}$, and $\sum_{K \in \mathcal{K}} p(K) = 1$;
3. r is a response function for (Q, \mathcal{K}, p) —that is, $r : 2^Q \times \mathcal{K} \rightarrow [0, 1]$, $(R, K) \mapsto r(R, K)$, with $r(R, K) \geq 0$ for any $R \in 2^Q$ and $K \in \mathcal{K}$, and $\sum_{R \in 2^Q} r(R, K) = 1$ for any $K \in \mathcal{K}$ (2^Q : power set of Q);
4. r satisfies local independence—that is, for any $R \in 2^Q$ and $K \in \mathcal{K}$,

$$r(R, K) = \left\{ \left[\prod_{q \in K \setminus R} \beta_q \right] \cdot \left[\prod_{q \in K \cap R} (1 - \beta_q) \right] \cdot \left[\prod_{q \in R \setminus K} \eta_q \right] \cdot \left[\prod_{q \in Q \setminus (R \cup K)} (1 - \eta_q) \right] \right\},$$

with constants $\beta_q, \eta_q \in [0, 1[$ for each $q \in Q$, respectively called careless error and lucky guess probabilities at q .

To each knowledge state $K \in \mathcal{K}$ is attached a probability $p(K)$ measuring the likelihood that an examinee is in state K (Point 2). For $R \in 2^Q$ and $K \in \mathcal{K}$, $r(R, K)$ specifies the conditional probability of response pattern R for an examinee in state K (Point 3). The item responses of an examinee are assumed to be independent given the knowledge state of the examinee. The response error probabilities β_q, η_q ($q \in Q$) are attached to the items and do not vary with the knowledge states (Point 4). Note that $K \setminus R := \{q \in Q : q \in K \text{ and } q \notin R\}$, $K \cap R := \{q \in Q : q \in K \text{ and } q \in R\}$, $R \setminus K := \{q \in Q : q \in R \text{ and } q \notin K\}$, and $Q \setminus (R \cup K) := \{q \in Q : q \notin R \text{ and } q \notin K\}$ form a partition of Q . The items in $K \setminus R$, $K \cap R$, $R \setminus K$, and $Q \setminus (R \cup K)$ are mastered but not solved (careless error), mastered and solved (no careless error), solved but not mastered (lucky guess), and not solved and not mastered (no lucky guess), respectively.

The BLIM allows expressing the (manifest) occurrence probabilities of the response patterns by means of the (latent) model parameters.

Corollary 1. Under the BLIM, the occurrence probabilities $\rho(R)$ of response patterns $R \in 2^Q$ are parameterized as

$$\rho(R) = \sum_{K \in \mathcal{K}} \left\{ \left[\prod_{q \in K \setminus R} \beta_q \right] \cdot \left[\prod_{q \in K \cap R} (1 - \beta_q) \right] \cdot \left[\prod_{q \in R \setminus K} \eta_q \right] \cdot \left[\prod_{q \in Q \setminus (R \cup K)} (1 - \eta_q) \right] \right\} p(K).$$

The parameters of the BLIM are $p(K)$ ($K \in \mathcal{K}$) and β_q, η_q ($q \in Q$). The number of independent model parameters is $2|Q| + (|\mathcal{K}| - 1)$. (For a set X , $|X|$ denotes the size of X .) Because the size of \mathcal{K} generally tends to be prohibitively large in practice, parameter estimation and model testing based on classical maximum likelihood methodology are not feasible in general (see, e.g., Ünlü, 2006). In KST, knowledge structures are built by qualitative or exploratory approaches; namely, by querying experts (e.g., Düntsch and Gediga, 1996; Koppen, 1993; Koppen and Doignon, 1990), from postulated psychological assumptions (e.g., Albert and Held, 1994; Albert, Schrepp, and Held, 1994; Düntsch and Gediga, 1995), or by numerical data analysis procedures (e.g., Sargin and Ünlü, 2008; Schrepp, 1999a,b, 2003). In this paper, we consider two KST exploratory data analysis methods in more detail, and demonstrate that these methods can be more effective when combined with the graphical mosaic plot technique. The mosaic plot provides an interesting new perspective on graphically displaying and deriving knowledge structures from data.

We close this section with a final remark. The connection of KST to other theories has been investigated in several publications. For instance, its connection to item response theory has been discussed by Stefanutti (2006) and Ünlü (2006, 2007). Schrepp (2005) and Ünlü (2006) have outlined KST's relationship to latent class (scaling) analysis (including Guttman scalogram analysis).

4 Simulations

All simulations and computations in this paper were realized using the R statistical computing environment (R Development Core Team, 2006, <http://www.r-project.org/>). The source files are freely available from the authors.

4.1 Data Generating Models

On a domain $Q = \{a, b, c, d, e\}$ of five dichotomous test items, we consider the three (antisymmetric) surmise relations \preceq_1 , \preceq_2 , and \preceq_3 depicted as Hasse diagrams in Figure 4.

According to Birkhoff's theorem (Theorem 1), the corresponding quasi ordinal knowledge spaces are, in respective order,

$$\begin{aligned} \mathcal{K}_{\preceq_1} &= \{\emptyset, \{e\}, \{d, e\}, \{c, d, e\}, \{b, c, d, e\}, Q\}, \\ \mathcal{K}_{\preceq_2} &= \{\emptyset, \{d\}, \{e\}, \{c, d\}, \{d, e\}, \{b, d, e\}, \{c, d, e\}, \{a, b, d, e\}, \{b, c, d, e\}, Q\}, \end{aligned}$$

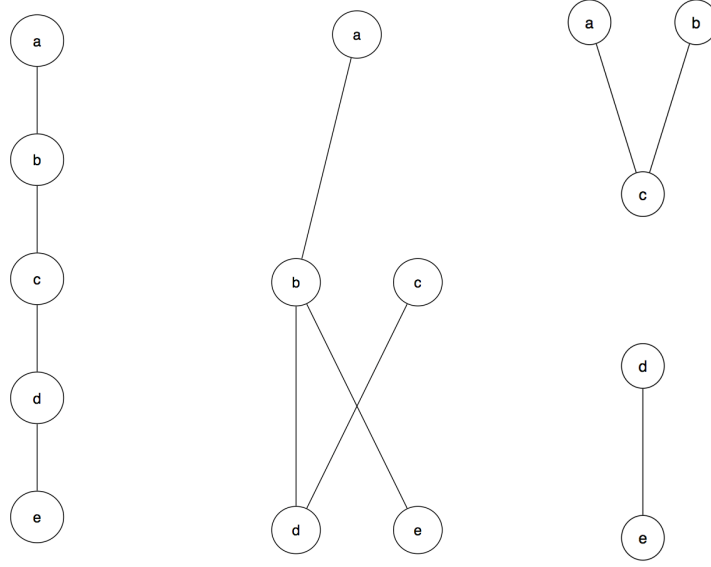


Fig. 4. Hasse diagrams of the three surmise relations \preceq_1 (left), \preceq_2 (middle), and \preceq_3 (right) on the domain $Q = \{a, b, c, d, e\}$. They consist of 15, 11, and 8 item pairs, respectively.

and

$$\mathcal{K}_{\preceq_3} = \{\emptyset, \{c\}, \{e\}, \{a, c\}, \{b, c\}, \{c, e\}, \{d, e\}, \{a, b, c\}, \{a, c, e\}, \{b, c, e\}, \{c, d, e\}, \{a, b, c, e\}, \{a, c, d, e\}, \{b, c, d, e\}, Q\}.$$

They consist of 6, 10, and 15 knowledge states, respectively.

Based on the BLIM (Section 3.3), we simulated data using each of these knowledge structures, endowed with the uniform probability distribution. We varied the sample size n , and a single response error rate $\beta_q = \eta_q = \tau$ ($q \in Q$). The aim is investigating the extent to which the underlying knowledge states can be visually recovered from mosaic plots of the simulated data.

4.2 Mosaic Plot Representation of the Knowledge Structures

The test items form an $|Q|$ -way dichotomy. Each knowledge state uniquely corresponds to a cell of this cross classification, which subsumes exactly the items mastered by the examinee. The tiles of the mosaic plot subsequently corresponding to the knowledge states of the knowledge structures \mathcal{K}_{\preceq_1} , \mathcal{K}_{\preceq_2} , and \mathcal{K}_{\preceq_3} are shown in Figure 5. Note that the two tiles in the upper left and lower right corners of the mosaic plot correspond to \emptyset and Q , respectively, which always are assumed to be knowledge states.

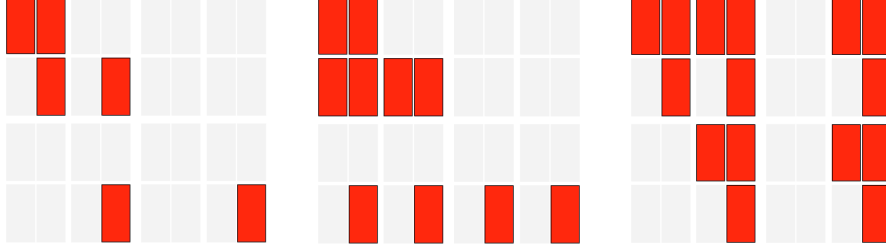


Fig. 5. Mosaic plot tiles corresponding to the knowledge states of the quasi ordinal knowledge spaces \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) (highlighted).

4.3 Results

First, we consider a ‘near to ideal’ situation; a large sample size $n = 1600$ and a small response error rate $\tau = 0.03$. Figure 6 shows the multiple barcharts views of the (three) datasets (one per each knowledge structure) simulated for these settings.

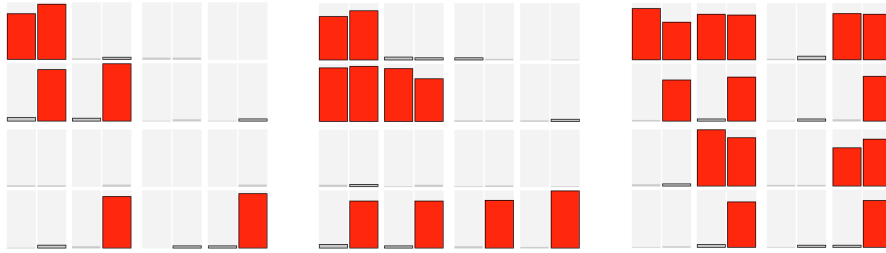


Fig. 6. Mosaic plots of the datasets simulated from \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) for $n = 1600$ and $\tau = 0.03$. The underlying knowledge states are highlighted.

The three multiple barcharts in Figure 6 give an unambiguous picture. For each of the knowledge structures, the tiles corresponding to the knowledge states clearly emerge, as compared to the ones that do not correspond to the states.

Next, we consider a more realistic situation; a small sample size $n = 100$ and a larger response error rate $\tau = 0.05$. Figure 7 shows the mosaic plots of the datasets simulated for these settings.

Apart from one or two exceptions for \mathcal{K}_{\preceq_3} , the mosaic plots still give a clear picture of the underlying knowledge states, albeit not as unambiguous as the previous ones. As compared to the plots in Figure 6, the tiles that do not correspond to the knowledge states now have larger heights but yet are discriminable from the tiles corresponding to the states. The more knowledge states appear, the more difficult it is to spot them. As shown in Figure 8

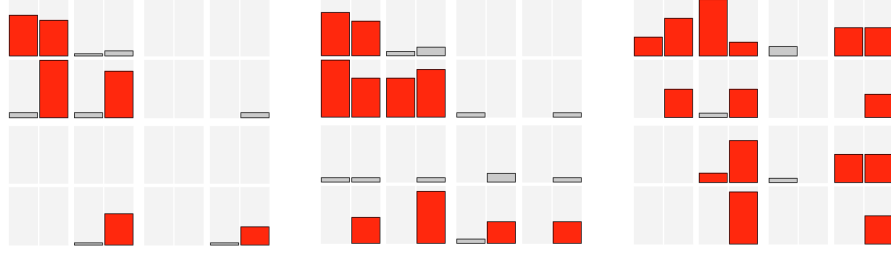


Fig. 7. Mosaic plots of the datasets simulated from \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) for $n = 100$ and $\tau = 0.05$. The underlying knowledge states are highlighted.

($n = 1600$, $\tau = 0.05$), the mosaic plots become better, and again unambiguous, when sample size increases.

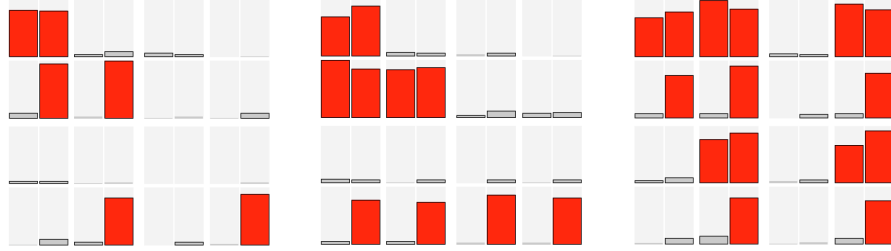


Fig. 8. Mosaic plots of the datasets simulated from \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) for $n = 1600$ and $\tau = 0.05$. The underlying knowledge states are highlighted.

Finally, we consider a ‘far from ideal’ situation; a small sample size $n = 100$ and a large response error rate $\tau = 0.20$. Figure 9 shows the multiple barcharts views of the datasets simulated for these settings.

Now, there are substantial ambiguities in the mosaic plot representation regarding the visual characterization of the underlying knowledge states. For \mathcal{K}_{\preceq_2} , for instance, four tiles of the (middle) plot in Figure 9 are critical. The knowledge state tile first (from left) in the second row (from top) of the mosaic plot has a height approximately the same as for the endmost two non-state tiles in that row. The non-state tile first in the fourth row even has a larger height than the latter state tile. Given the fact that the tiles in the upper left and lower right corners always correspond to knowledge states, the results worsen with increasing size of the underlying knowledge structure. However, this is a very difficult situation, in which, for instance, numerical data analysis methods for deriving knowledge structures also fail (see Section 6). Nevertheless, as shown in Figure 10 ($n = 1600$, $\tau = 0.20$), the mosaic plots again become better, and even acceptable, when sample size increases.

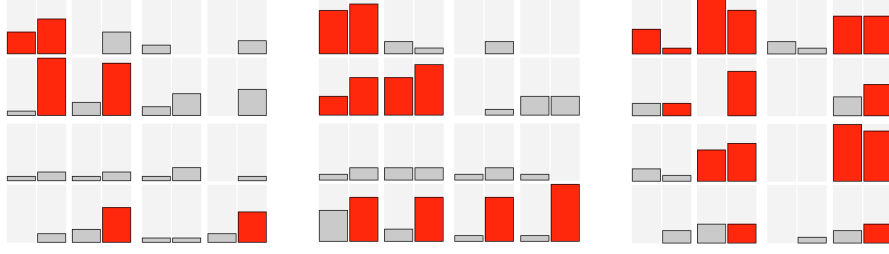


Fig. 9. Mosaic plots of the datasets simulated from \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) for $n = 100$ and $\tau = 0.20$. The underlying knowledge states are highlighted.

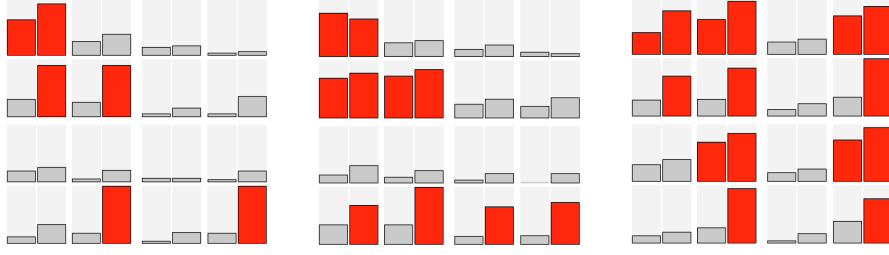


Fig. 10. Mosaic plots of the datasets simulated from \mathcal{K}_{\preceq_1} (left), \mathcal{K}_{\preceq_2} (middle), and \mathcal{K}_{\preceq_3} (right) for $n = 1600$ and $\tau = 0.20$. The underlying knowledge states are highlighted.

In summary: The more knowledge states were used, the more difficult it was to recover them visually in mosaic displays. Smaller response error rates implied more reliable graphical detection. With increasing sample size the results became even better.

5 Application to Empirical Data

In this section, the mosaic plot is applied for detecting knowledge states in an empirical dataset, which is from the Programme for International Student Assessment (PISA; <http://www.pisa.oecd.org/>). Static plots do not allow interacting with graphics. User interaction (e.g., Theus and Urbanek, 2008; Unwin et al., 2006) is seen to be helpful in exploring these assessment data.

5.1 PISA Dataset

We analyze part of the 2003 PISA data which consists of 340 German students answering five questions on mathematical literacy; $Q = \{a, b, c, d, e\}$. The dataset (available from the authors) consists of 0/1 scores; a correct answer is scored 1, and an incorrect answer 0. These items form a Rasch scale. That is,

the (dichotomous) one-parameter logistic item response theory model (Fischer and Molenaar, 1995) fits the data very well (goodness-of-fit and item fit). Under this model, the following item difficulties for the five questions are estimated: -2.07 (Item a), -1.22 (Item b), -0.04 (Item c), 1.44 (Item d), and 2.18 (Item e). Hence, the items most likely form a chain, which is considered as the underlying surmise relation in the subsequent analyses (see Figure 11).

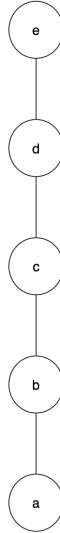


Fig. 11. Hasse diagram of the Rasch scale of five assessment items on the domain $Q = \{a, b, c, d, e\}$. From bottom to top, sorted according to increasing difficulty. Assumed to underlie the PISA dataset.

The corresponding quasi ordinal knowledge space (Theorem 1) consists of the states \emptyset , $\{a\}$, $\{a, b\}$, $\{a, b, c\}$, $\{a, b, c, d\}$, and Q .

5.2 Results

Figure 12 shows the multiple barcharts view of the PISA dataset.

The multiple barcharts in Figure 12 give a satisfactory picture. The two tiles in the upper left and lower right corners of the mosaic plot correspond to the knowledge states \emptyset and Q . The tiles representing the remaining states reasonably emerge, as compared to the ones that do not correspond to the states.

In Figure 13, the three (most likely) non-state tiles with relatively large heights are queried.

Queried information, especially when utilized in combination with linked plots, may prove valuable in studying specific aspects of a phenomenon. For

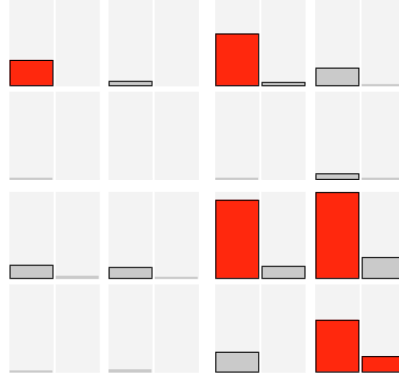


Fig. 12. Mosaic plot of the PISA dataset. The underlying knowledge states are highlighted.

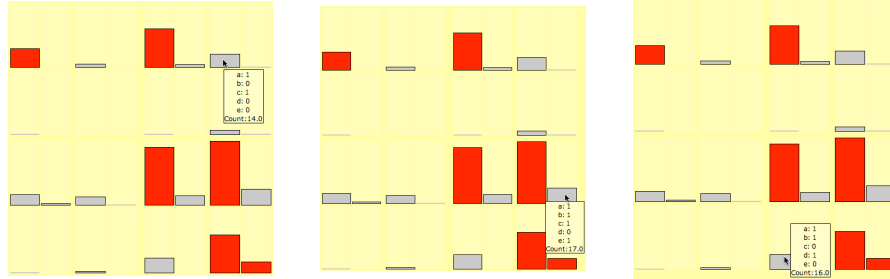


Fig. 13. Standard querying gives basic information about the three non-state tiles with relatively large heights. The underlying knowledge states are highlighted.

instance, consider the middle plot of Figure 13. If the data are open format test data such that guessing effects are nearly eliminated by appropriate item formulation, the majority of the cases landing in the queried tile most probably committed a careless error on Item d . This could furthermore explain why, compared to that non-state tile, there are fewer cases contained in the lower right corner tile, which corresponds to the knowledge state Q .

Other useful interactive visualization techniques are, for instance, aggregation and linking. Figure 14 shows a mosaic plot for the Items a , b , c , and e of the PISA dataset, aggregating over Item d , and a barchart for Item d . The examinees solving that particular item are selected and highlighted.

Interactive queries give following information. There are 79 examinees solving Item d . Almost all of these cases fall into the three tiles of the aggregated mosaic plot corresponding to the induced states $\{a, b\}$ (third row, third column), $\{a, b, c\}$ (third row, fourth column), and $\{a, b, c, e\}$ (fourth row, fourth column). The absolute / relative frequencies are, in respective order, 16 / 20.78%, 40 / 37.38%, and 12 / 41.38%. Note that these proportions can also be seen from the mosaic plot. The relative frequencies are estimates of the

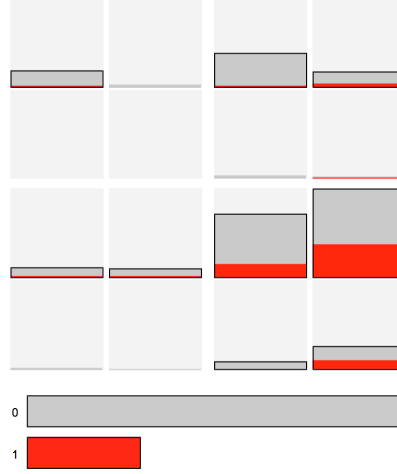


Fig. 14. Aggregation and linking. Mosaic plot (top) for the Items a , b , c , and e of the PISA dataset, aggregating over Item d . Barchart (bottom) for Item d . Examinees solving Item d are selected and highlighted.

conditional probabilities that Item d is solved given any of the former induced states. The obtained figures support the underlying chain hierarchy among the five test items. Item d is a successor to Item c , and a predecessor to Item e . In other words, graphical considerations of such a type may help in classifying (e.g., new) test items appropriately.

6 Numerical Data Analysis Methods

In this section, two KST numerical data analysis methods for deriving surmise relations on sets of dichotomous items are considered. These methods are shown to be more effective when applied in combination with the mosaic plot.

6.1 Inductive Item Tree Analysis

The two inductive item tree analysis (IITA) algorithms considered in this paper are the original algorithm by Schrepp (1999a, 2003, 2006), and the corrected version by Sargin and Ünlü (2008) and Ünlü and Sargin (2008b). In both methods, competing surmise relations are generated, and a fit measure is computed for every relation in order to find the one that fits the data best.

The first step, the same for both algorithms, is the inductive generation of surmise relations. For two items i and j , $b_{ij} = |\{R \in D | i \notin R \wedge j \in R\}|$ is the number of counterexamples, that is, the number of observed response patterns in dataset D contradicting $i \preceq j$ (mastering j implies mastering i). Based on these values, binary relations \preceq_L for $L = 0, \dots, n$ are defined.

Let $i \preceq_0 j \Leftrightarrow b_{ij} = 0$. The relation \preceq_0 is transitive, and based on that, all the other transitive relations \preceq_L are constructed inductively. Assume \preceq_L is a transitive relation. Define the set $S_{L+1} = \{(i, j) | b_{ij} \leq L+1 \wedge i \not\preceq_L j\}$. This set consists of all item pairs that are not already contained in the relation \preceq_L and have at most $L+1$ counterexamples. From these item pairs those are excluded that cause an intransitivity in $\preceq_L \cup S_{L+1}$, and the remaining item pairs are referred to as $S_{L+1}^{(1)}$. This process continues iteratively, say k times, until no intransitivity is caused anymore. The generated relation $\preceq_{L+1} := \preceq_L \cup S_{L+1}^{(k)}$ is then transitive by construction. Because \preceq_0 is reflexive, all generated relations are. Hence \preceq_L is a quasi order for every $L = 0, \dots, n$.

Besides the construction of the surmise relations, in a second step, it is very important to find that surmise relation which fits the data best. The idea is to estimate the numbers of counterexamples for each surmise relation, and to find, over all competing surmise relations, the minimum value for the discrepancy between the observed and expected numbers of counterexamples.

Let $p_i := |\{R \in D | i \in R\}|/n$ be the relative solution frequency of an item i . A violation of an underlying dependency is only possible due to random errors. To compute the expected numbers of counterexamples, b_{ij}^* , error probabilities are needed. In both algorithms, the error probabilities are assumed to be equal for all items, and that single error rate is estimated by

$$\gamma_L := \frac{\sum \{b_{ij} / (p_j n) | i \preceq_L j \wedge i \neq j\}}{(|\preceq_L| - |Q|)},$$

where $|\preceq_L| - |Q|$ is the number of non-reflexive item pairs in \preceq_L . Under every relation, the algorithms compute the expected numbers of counterexamples for each (non-reflexive) item pair. First we present the estimates used in the original IITA algorithm. If the relation \preceq_L provides a dependency $i \preceq_L j$, the expected number of counterexamples is computed by $b_{ij}^* = p_j \gamma_L n$. If $(i, j) \notin \preceq_L$, no relationship between the two items is assumed, and $b_{ij}^* = (1 - p_i) p_j n (1 - \gamma_L)$. For a discussion and criticism on these estimates, see Sargin and Ünlü (2008). In the corrected IITA algorithm, the following estimates are used. If $(i, j) \in \preceq_L$, set $b_{ij}^* = p_j \gamma_L n$. If $(i, j) \notin \preceq_L$ and $(j, i) \notin \preceq_L$, set $b_{ij}^* = (1 - p_i) p_j n$. If $(i, j) \notin \preceq_L$ and $(j, i) \in \preceq_L$, set $b_{ij}^* = (p_j - p_i + p_i \gamma_L) n$. Motivation for and derivation of these estimates can be found in Sargin and Ünlü (2008).

The fit of each relation \preceq_L to the dataset D is quantified using the *diff* coefficients

$$diff(\preceq_L, D) = \sum_{i \neq j} \frac{(b_{ij} - b_{ij}^*)^2}{|Q|(|Q| - 1)},$$

where corresponding estimates b_{ij}^* are used. They give the average sums of the quadratic differences between the observed and expected numbers of counterexamples under the relation \preceq_L . The smaller the *diff* values are, the better is the fit of the relation to the data. Therefore, the IITA algorithms look

for the smallest values of the *diff* coefficients and return the corresponding surmise relations.

6.2 Results

We reconsider the dataset simulated from \mathcal{K}_{\preceq_3} for a sample size $n = 1600$ and a response error rate $\tau = 0.20$, visualized in Figure 10 (right plot) of Section 4.3. Analyses of this dataset using the original and corrected IITA algorithms give the results reported in Figure 15.

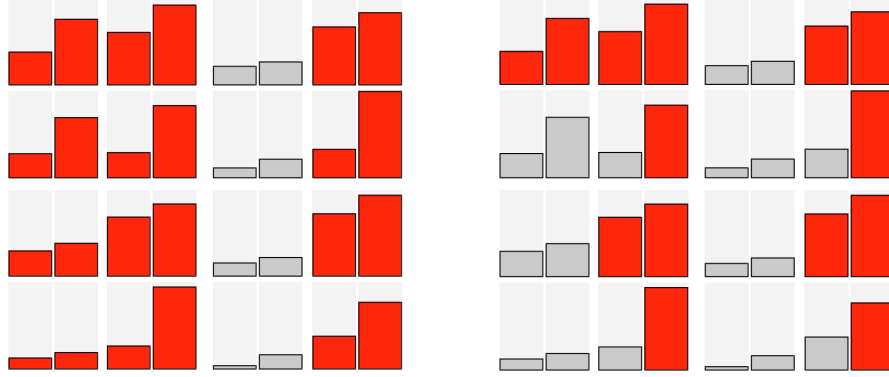


Fig. 15. Mosaic plot of the dataset simulated from \mathcal{K}_{\preceq_3} for $n = 1600$ and $\tau = 0.20$ (cf. Figure 10). The knowledge states obtained for this dataset under the original (left) and corrected (right) IITA algorithms are highlighted.

The original IITA algorithm is clearly outperformed by the corrected. The former erroneously includes nine non-states (e.g., fourth row, first column), the latter misses only one state (second row, second column). This confirms the results obtained in Sargin and Ünlü (2008) and Ünlü and Sargin (2008b). More importantly, seeing the mosaic plot the knowledge state missed by the corrected IITA version could be detected visually and be added subsequently, giving the true knowledge structure underlying the dataset. The mosaic plot clearly substantiates the implausibility of the knowledge structure returned by the original IITA algorithm (e.g., first two non-state tiles in fourth row).

Figure 16 shows the knowledge structure resulting from an original IITA analysis of a dataset simulated from \mathcal{K}_{\preceq_1} for a sample size $n = 400$ and a response error rate $\tau = 0.03$. Note that this dataset has not been used before; however, see Figure 6 (left plot).

The multiple barcharts in Figure 16 give an unambiguous picture. Indeed, they provide valuable information. The tile corresponding to the knowledge state missed by the original IITA algorithm (second row, second column) clearly emerges, as compared to the non-state tiles. This knowledge state

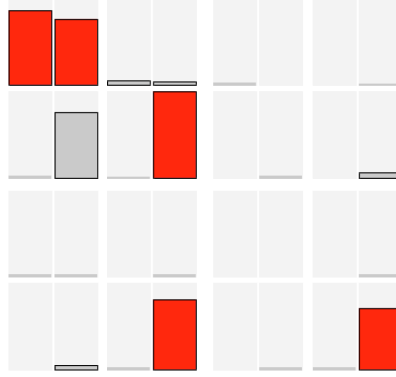


Fig. 16. Mosaic plot of the dataset simulated from $\mathcal{K}_{\leq 1}$ for $n = 400$ and $\tau = 0.03$ (cf. Figure 6). The knowledge states obtained for this dataset under the original IITA algorithm are highlighted.

could be added subsequently, giving the true knowledge structure underlying the simulated dataset. The corrected IITA algorithm applied to this dataset yields $\mathcal{K}_{\leq 1}$.

Next, we reconsider the PISA dataset (Section 5), visualized in Figure 12. Analyzing this empirical dataset using the original and corrected IITA algorithms gives the results reported in Figure 17.

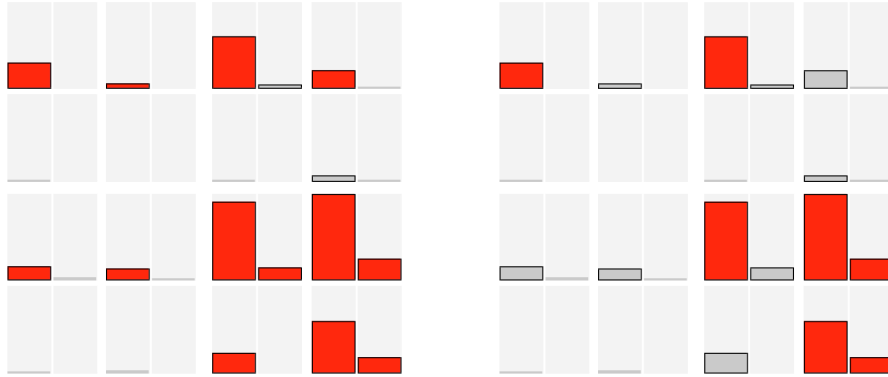


Fig. 17. Mosaic plot of the PISA dataset (cf. Figure 12). The knowledge states obtained for this dataset under the original (left) and corrected (right) IITA algorithms are highlighted.

The original and corrected IITA algorithms both detect all the underlying knowledge states. The original IITA algorithm additionally includes seven non-states; for instance, the non-state represented by the tile in the first row and third column. Using the mosaic plot, this tile would certainly be discarded.

Note that the tiles in the upper left and lower right corners correspond to the knowledge states \emptyset and Q . The corrected IITA algorithm, on the other hand, only includes one non-state. Yet the tile representing this non-state (third row, last column) has a relatively large height (cf. also the remarks to Figure 13). The inclusion of this non-state could probably be avoided using the mosaic plot. The empty set and the domain are always states. The heights of the remaining state tiles are clearly larger than the height of that non-state tile. Once again, the original IITA algorithm is outperformed by the corrected IITA algorithm.

7 Conclusion

We have investigated the scope of the mosaic plot for visually inspecting latent knowledge states in discrete multivariate response data in KST. We have seen that this graphing method reveals the underlying knowledge structure very well for smaller response error rates. The results become even better with increasing sample size. The more knowledge states occur, the more difficult it is to recover them. The mosaic plot has also been satisfactorily applied to the PISA dataset. Since numerical data analysis procedures in KST may produce implausible findings, a graphical approach such as the present one can provide valuable information with which to compare the results obtained from numerical methods. Graphics, especially when implemented in interactive software, are particularly valuable for exploratory work (e.g., Theus and Urbanek, 2008). They complement rather than replace confirmatory statistical approaches.

Future research on graphics in KST could include visualizing combinatorial properties of a knowledge structure. The simplest such property is closedness under union. A more advanced property is wellgradedness (item-wise gradual learning), which leads to the pedagogically important concept of a learning space. Building substructures of a knowledge structure on subsets of the item set is also important. Knowledge structures may be too large to work with conveniently, or only a specific part of the domain may be needed. This can graphically be realized, for instance, through interactive mosaic plots. Since combinatorial structure—not data—is to be visualized, the mosaic plot representation of a knowledge structure introduced in Figure 5 is adequate. Interactively aggregating over items in such a mosaic plot representation allows easily visualizing substructures. In particular, the aggregated mosaic plot then shows how many states of the parent structure induce the same traces on the subdomain. Selecting a trace state of the substructure, in turn, allows visually displaying the corresponding knowledge states of the parent structure.

To conclude, the mosaic plot can become quickly intractable, space and display resolution are limiting factors (e.g., Hofmann, 2008). Nevertheless, it represents a promising novel perspective on empirically evaluating knowledge structure models in KST.

References

- D. Albert and T. Held. Establishing knowledge spaces by systematical problem construction. In D. Albert, editor, *Knowledge Structures*, pages 78–112. Springer-Verlag, Berlin and Heidelberg, 1994.
- D. Albert and J. Lukas, editors. *Knowledge Spaces: Theories, Empirical Research, and Applications*. Lawrence Erlbaum Associates, Mahwah, NJ, 1999.
- D. Albert, M. Schrepp, and T. Held. Construction of knowledge spaces for problem solving in chess. In G.H. Fischer and D. Laming, editors, *Contributions to Mathematical Psychology, Psychometrics, and Methodology*, pages 123–135. Springer-Verlag, New York, 1994.
- G. Birkhoff. Rings of sets. *Duke Mathematical Journal*, 3:443–454, 1937.
- J.M. Chambers, W.S. Cleveland, B. Kleiner, and P.A. Tukey. *Graphical Methods for Data Analysis*. Duxbury Press, Boston, MA, 1983.
- C.-H. Chen, W. Härdle, and A.R. Unwin, editors. *Handbook of Data Visualization*. Springer-Verlag, Berlin, 2008.
- W.S. Cleveland. *The Elements of Graphing Data*. Hobart Press, Summit, NJ, revised edition, 1994.
- D. Cook and D. Swayne. *Interactive and Dynamic Graphics for Data Analysis*. Springer-Verlag, New York, 2007.
- J.-P. Doignon and J.-Cl. Falmagne. Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23:175–196, 1985.
- J.-P. Doignon and J.-Cl. Falmagne. Knowledge assessment: A set theoretical framework. In *Beiträge zur Begriffsanalyse: Vorträge der Arbeitstagung Begriffsanalyse, Darmstadt, 1986*, pages 129–140, Mannheim, Germany, 1987. B.I. Wissenschaftsverlag.
- J.-P. Doignon and J.-Cl. Falmagne. *Knowledge Spaces*. Springer-Verlag, Berlin, Heidelberg, and New York, 1999.
- I. Düntsch and G. Gediga. Skills and knowledge structures. *British Journal of Mathematical and Statistical Psychology*, 48:9–27, 1995.
- I. Düntsch and G. Gediga. On query procedures to build knowledge structures. *Journal of Mathematical Psychology*, 40:160–168, 1996.
- J.W. Emerson. Mosaic displays in S-Plus: A general implementation and a case study. *Statistical Computing & Statistical Graphics Newsletter*, 9: 17–23, 1998.
- J.-Cl. Falmagne. Probabilistic knowledge spaces: A review. In F.S. Roberts, editor, *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences*, volume 17, pages 283–303. Springer-Verlag, New York, 1989.
- J.-Cl. Falmagne. Projections and symmetric expansions of a learning space. Electronic preprint 0803.0575v1, arXiv.org, 2008.
- J.-Cl. Falmagne and J.-P. Doignon. A class of stochastic procedures for the assessment of knowledge. *British Journal of Mathematical and Statistical Psychology*, 41:1–23, 1988a.

- J.-Cl. Falmagne and J.-P. Doignon. A Markovian procedure for assessing the state of a system. *Journal of Mathematical Psychology*, 32:232–258, 1988b.
- J.-Cl. Falmagne, M. Koppen, M. Villano, J.-P. Doignon, and L. Johannesen. Introduction to knowledge spaces: How to build, test and search them. *Psychological Review*, 97:201–224, 1990.
- J.-Cl. Falmagne, E. Cosyn, J.-P. Doignon, and N. Thiéry. The assessment of knowledge, in theory and in practice. In *Formal Concept Analysis, 4th International Conference, ICFCA 2006, Dresden, Germany, Lecture Notes in Artificial Intelligence*, pages 61–79, Berlin, Heidelberg, and New York, 2006. Springer-Verlag.
- G.H. Fischer and I.W. Molenaar, editors. *Rasch Models: Foundations, Recent Developments, and Applications*. Springer-Verlag, New York, 1995.
- M. Friendly. Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89:190–200, 1994.
- M. Friendly. Conceptual and visual models for categorical data. *American Statistician*, 49:153–160, 1995.
- M. Friendly. Extending mosaic displays: Marginal, conditional, and partial views of categorical data. *Journal of Computational and Graphical Statistics*, 8:373–395, 1999.
- M. Friendly. *Visualizing Categorical Data*. SAS Institute Inc., Cary, NC, 2000.
- J.A. Hartigan and B. Kleiner. Mosaics for contingency tables. In *Computer Science and Statistics: Proceedings of the Thirteenth Symposium on the Interface*, pages 268–273, New York, 1981. Springer-Verlag.
- J.A. Hartigan and B. Kleiner. A mosaic of television ratings. *American Statistician*, 38:32–35, 1984.
- H. Hofmann. Simpson on board the titanic? interactive methods for dealing with multivariate categorical data. *Statistical Computing & Statistical Graphics Newsletter*, 9:16–19, 1998.
- H. Hofmann. Exploring categorical data: Interactive mosaic plots. *Metrika*, 51:11–26, 2000.
- H. Hofmann. Constructing and reading mosaicplots. *Computational Statistics & Data Analysis*, 43:565–580, 2003.
- H. Hofmann. Mosaic plots and their variants. In C.-H. Chen, W. Härdle, and A.R. Unwin, editors, *Handbook of Data Visualization*, pages 617–642. Springer-Verlag, Berlin, 2008.
- H. Hofmann and A.F.X. Wilhelm. Visual comparison of association rules. *Computational Statistics*, 16:399–415, 2001.
- H. Hofmann, A.P.J.M. Siebes, and A.F.X. Wilhelm. Visualizing association rules with interactive mosaic plots. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 227–235, New York, 2000. Association for Computing Machinery.
- J. Hummel. Linked bar charts: Analysing categorical data graphically. *Computational Statistics*, 11:23–33, 1996.
- M. Koppen. Extracting human expertise for constructing knowledge spaces: An algorithm. *Journal of Mathematical Psychology*, 37:1–20, 1993.

- M. Koppen and J.-P. Doignon. How to build a knowledge space by querying an expert. *Journal of Mathematical Psychology*, 34:311–331, 1990.
- P. Murrell. *R Graphics*. Chapman & Hall/CRC, Boca Raton, FL, 2005.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2006. ISBN 3-900051-07-0.
- A. Sargin and A. Ünlü. Inductive item tree analysis: Corrections, improvements, and comparisons. Manuscript submitted for publication, 2008.
- M. Schrepp. On the empirical construction of implications between bi-valued test items. *Mathematical Social Sciences*, 38:361–375, 1999a.
- M. Schrepp. Extracting knowledge structures from observed data. *British Journal of Mathematical and Statistical Psychology*, 52:213–224, 1999b.
- M. Schrepp. A method for the analysis of hierarchical dependencies between items of a questionnaire. *Methods of Psychological Research*, 19:43–79, 2003.
- M. Schrepp. About the connection between knowledge structures and latent class models. *Methodology*, 1:93–103, 2005.
- M. Schrepp. ITA 2.0: A program for classical and inductive item tree analysis. *Journal of Statistical Software*, 16(10), 2006.
- L. Stefanutti. A logistic approach to knowledge structures. *Journal of Mathematical Psychology*, 50:545–561, 2006.
- M. Theus. Interactive data visualization using Mondrian. *Journal of Statistical Software*, 7(11), 2002.
- M. Theus and S. Urbanek. *Interactive Graphics for Data Analysis*. CRC Press, London, 2008.
- A. Ünlü. Estimation of careless error and lucky guess probabilities for dichotomous test items: A psychometric application of a biometric latent class model with random effects. *Journal of Mathematical Psychology*, 50:309–328, 2006.
- A. Ünlü. Nonparametric item response theory axioms and properties under nonlinearity and their exemplification with knowledge space theory. *Journal of Mathematical Psychology*, 51:383–400, 2007.
- A. Ünlü and A. Sargin. Interactive visualization of assessment data: The software package Mondrian. *Applied Psychological Measurement*, 2008a. In press.
- A. Ünlü and A. Sargin. Maximum likelihood methodology for *diff* fit measures for quasi orders. Manuscript submitted for publication, 2008b.
- A.R. Unwin, G. Hawkins, H. Hofmann, and B. Siegl. Manet – Extensions to interactive statistical graphics for missing values. In *New Techniques and Technologies for Statistics II*, pages 247–259, Amsterdam, 1997. IOS Press.
- A.R. Unwin, M. Theus, and H. Hofmann. *Graphics of Large Datasets*. Springer-Verlag, New York, 2006.
- S. Urbanek. Different ways to see a tree - KLIMT. In *15th Conference on Computational Statistics, COMPSTAT*, pages 303–308, Heidelberg, Germany, 2002. Physica.

- S. Urbanek and M. Theus. iPlots: High interaction graphics for R. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, DSC 2003*, Vienna, Austria, 2003. ISSN 1609-395X.
- L. Wilkinson. *The Grammar of Graphics*. Springer-Verlag, New York, 2nd edition, 2005.
- F. Young, P. Valero-Mora, and M. Friendly. *Visual Statistics: Seeing Data with Dynamic Interactive Graphics*. Wiley, Hoboken, NJ, 2006.